DOCUMENT RESUME

ED 036 766 AC 006 572

AUTHOR	Ilyin, Donna
TITLE	Developing a Placement Test for Adults in English-Second-Language Programs in California.
PUB DATE	Feb 70
NOTE	30p.; Paper presented at the Adult Education Research Conference, Minneapolis, Minnesota, February 27-28, 1970
EDRS PRICE	EDRS Price MF-$0.25 HC-$1.60
DESCRIPTORS	*English (Second Language), *Instructional Program Divisions, *Student Placement, *Test Construction

ABSTRACT

This paper describes the development of the English-Second-Language Placement Test (EPT) 100-200-300, which places adult students into the first three levels of ESL classes, and discusses work done on EPT 400-500-600, an experimental test to place students in the last three levels of classes. A structured test, EPT 100-200-300 tests ability to read items already practiced orally. One-page pretests were standardized on students at Alemany Adult School in San Francisco. Subsequent work was done under a Federal Grant to establish norms and to make a gains study on small city day schools and on night schools in urban and migrant worker areas. Scores of adult non-academic students were compared with those of college students enrolled in English classes for the foreign born at San Francisco State College. The new standardized tests are called EPT 100-200-300 forms A and B; each contains 50 items, takes a half hour, and provides a placement table. EPT 400-500-600 is still experimental but students successfully passing this test should be able to compete in Adult High Schools or College Foreign Born programs. (author/eb)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED036766

DEVELOPING A PLACEMENT TEST FOR ADULTS IN ENGLISH - SECOND - LANGUAGE PROGRAMS IN CALI[illegible]

by
Donna ILYIN

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY Donna Ilyin
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

I want to describe my development of EPT 100-200-300, a standardized test which places adult second-language students into the first three levels of English-Second Language classes and to discuss work done on EPT 400-500-600, an experimental test I developed with Mrs. Jeanette Anderson Best and Mrs. Virginia Biagi to place students in the last three levels of English-Second-Language classes. EPT (English-Second Language Placement Test) 100-200-300 has an overall reliability of .96 and form correlation of .93. EPT 400-500-600 is still experimental.

Ten years ago when I first began teaching English as a second language in an adult program in San Francisco, I was surprised at the utter chaos that existed in our classes. I wondered why we had so many students of such diversified abilities and backgrounds in each class. It seemed only logical to me as well as to many other teachers that we needed a test that would be able to place students better than the hit or miss interview in the main office.

Of course at the beginning, I never realized that we would have to develop our own tests. I naturally assumed that we only had to find placement tests for English-Second-Language classes and just administer them. We soon discovered, however, that tests for adult students were either directed at first-language learners, to place students in grade levels or at college oriented foreign born students. Until our

AC006572

students reached the upper levels in our school, they were unable to take the tests. The majority of our students were in the lower levels.

I developed the test and supervised the testing work of Mrs. Best and Mrs. Biagi without knowing that classroom teachers usually didn't make standardized tests. This is considered to be the sovereign domain of testing services and University professors highly trained in psycholinguistics and testing. However, their interests apparently lie in other fields that may be lush and green, but that are barely visible from classrooms of the non-academic adult English-Second-Language learners. We were on the spot and needed a classifying instrument which could be administered to large groups of students in a short time and which could be graded by clerical personnel.

I believe that when a student is placed at a level where he can realistically handle material and feel that he is succeeding and learning, he will stay in a class and will become more proficient in English. A teacher can do a better job of structuring lessons, setting realistic objectives, and providing a better learning climate when his students are more or less at the same level of ability. At the beginning levels most of the work is oral and it is difficult to provide interesting and challenging work when the students' range of abilities is too great. Until teachers get good sequentially developed programmed materials designed to meet the individual language problems of each student and each first language background, no machines or small grouping will adequately solve the present problem found in heterogeneous classes of wide ranges of ability.

My rationale for making a structure test was that most second-language learning materials are built around structural patterns which

are sequentially presented. Vocabulary is limited in early levels of ESL instruction. Reading and writing are usually confined to copying and manipulating oral patterns already mastered.

I made an objective multiple choice test in a written form because it was easier to make, easier to take, and because we needed a short test with simple directions that could be administered and graded quickly. J. B. Carroll reports that short tests of integrated language performance have been found to be just as valid if the items are complex and cover a wide sample of linguistic competence as a test that has each item contrived to cover one and only one specific feature.[1]

1 J.B. Carroll, "The Psychology of Language Testing", Language Testing Symposium, A Psycho-linguistic Apprcach, edited by Alan Davies, Oxford University Press, 1968, page 56.

Carroll also says that we usually test productive and receptive skills separately because they are less likely to be highly correlated. I did not do this. There has been very little research to date on second-language skill testing and correlation of productive and receptive skills. In my field there have been no standarized tests simple enough to administer that measure these skills independently. Like most California non-academic adult students taught by trained ESL teachers who use second-language materials, our students can read patterns that they have been overdrilled on in class even though no formal instruction has been given in reading.[2]

2 There seems to be a great deal of misunderstanding and confusion about a student's ability to read patterns he has learned orally. Wilga Rivers in her book, Teaching Foreign Language Skills, University of Chicago Press, 1968, discusses this problem thoroughly on pages 217-220. Charles Fries in Teaching and Learning English as a Foreign Language, University

of Michigan Press, 1945, referred to the problem in a footnote on page 136. See also Roy Steeves and Patricia H. Cabrera, Handbook for Teachers of English as a Second Language, Americanization-Literacy, California State Department of Education, 1969, pages 18, 21 and 24.

While I made an objective multiple choice test that required a student to look at a printed form of a pattern he had practiced in class, and choose the pattern he thought was correct, my distractors (wrong answers) differed from those in the usual English multiple choice tests designed for native born speakers of English. For six years, I had kept a card file of errors my students made in patterns. My distractors came from those errors made before students learned a correct pattern. For this reason my test (although in a written form) tends to test what students think is correct from their oral practice and conversation. (A statistician in the testing office at San Francisco State College remarked that those errors were similar to those deaf children made.)

When I began teaching ten years ago, our school was a full-time adult day school attached to the adult high school program. At that time we had five levels of instruction with six hours of classes arranged into periods of 2 hours each. Students ranged in age from 18 - 80, had a wide range of education, and represented about 60 different language groups. They usually lived with people who spoke their native or first language and may or may not have had some kind of exposure to English. Many were not highly motivated, committed students and were not anxious and willing to learn English actively.

Teachers chose their own text books and decided when a student was ready for a higher level, which could happen anytime during the

semester. No grades were given and students enrolled in class and left at will at all points during the class semester. There was much shopping around. Some students became attached to one teacher and stayed with him a number of terms. If teachers did not maintain a daily attendance of 14, their classes were dropped and they often lost their jobs.

Many of the teachers were traditional English High School teachers and attempted to teach our second-language learners with high school materials for native born students or with children's readers and spellers. Some of the teachers were speech or high school foreign language teachers who used traditional materials designed years ago for immigrants or second-language learners. A few were trained or being trained in the audio-lingual system and adapted second-language materials used in overseas programs or in courses for foreign born students enrolled in American colleges.

At first no one could agree on what was to be taught at which level. Many of us held our classes on our personality, showmanship, sociability, and interest in helping the students learn English. Some of us met and discussed what we could do to realistically help our students learn English. Many of us enrolled in linguistic courses.

William Tresnon, one of our registrars at that time and now the principal of our school, felt the same concern we did and asked us to form committees to solve two basic problems:

1. What to teach at each level.
2. How to place new students in the levels.

From the committee groups, Mr. Tresnon along with James Norris, now a registrar at Pacific Heights Adult School in San Francisco, wrote a curriculum guide, A Sequential Course of Study in English for the

Foreign Born , San Francisco Unified School District, 1967. This helped those of us attempting to make a placement test.

Our first efforts were voluntary. Three of us worked on a committee to develop an achievement test which we administered to classes at the first three levels. We then regrouped the students and the teachers decided which text books would be used at which levels. This test had taken a long time to develop and to grade and it became compromised. A teacher at another school obtained a copy of it and used it as a lesson review. Copies of the test floated through our school. Our committee lost interest in developing another test until we could have some assurance the test's security would be protected. We also moved to another building without a large room where we could administer the test and control the examination.

Some of us continued to make our own tests so that we could change students erroneously placed in our classes as quickly as possible. These tests were cumbersome to grade. Once again, we tried to administer an achievement type test in our lower level classes, but met with resistance from some teachers who refused to proctor adequately, and from others who feared some kind of negative judgment of their classes.

In March, 1966, Mr. Dalton Howatt, the adult coordinator of the San Francisco Unified School District gave me some released time to develop an objective instrument to place students in our classes. I was also allowed to administer all tests and maintain necessary test security. At first, I hoped to develop listening and writing tests as well as grammar structure tests, but finally I limited myself to the grammar structure tests. We decided to make pretests (with

two forms) for each level. I had a pamphlet written by David Harris which became a bible of procedure.[3]

---

3 David P. Harris, English Testing Guidebook, Part II, Fundamentals of Test Construction and Interpretation, American University Language Center for the International Cooperation Administration, June, 1961.

Mr. Robert Breckenridge, one of our teachers who had worked on testing in the Human Resources Research Center, U.S.A.F. answered many of my questions.

Some teachers at the school contributed their achievement tests and volunteered to grade papers. I analyzed all books used at each level and made sample tests from which I obtained student's errors. From these I made lists of from 200-300 suggested items for an objective test that covered patterns taught at each level. Again teachers at our school voluntarily analyzed the items and made additions or suggestions.

Then for each level, I made two final sample objective tests containing from 50-75 items each, which I gave to students. Next by item analysis, I constructed for each level, two one-page pretests that could be administered to large groups of students at one time and that could be graded in less than a minute per test. Each page of legal size mimeograph paper contained 30 three-choice items.

Students read each item, considered their choices (A.B.C.) and put an X on the alphabet letter proceeding the item they thought correct. Before beginning the test, students answered trial questions to determine if they understood how to take the test. Not until everyone could follow the simple instructions did we begin. Students then had 15 minutes to do 30 items. Sample trial items are as follows:

1. A. I is here.
   B. I am here.
   C. I are here.

2. What is that?

   A. It is a book.
   B. He is a book.
   C. She is a book.

We were careful to select items that represented various grammar problems and that were statistically the most discriminating and reliable. By June, 1968, I had developed eight pretests, (two forms for each of our first four levels) and had a bank of 300 suggested items for our level five. Reliability on the tests were from .74 to .84 and form correlations at each level were from .76 to .86. Dr. Henry Clay Lindgren from San Francisco State College supervised the statistical work on my pretests. (See Attachment D for pretest statistics.)

By 1968 more of our teachers had had linguistic training; many had Masters Degrees in teaching English as a Second Language. We were beginning to have a good sound ESL program and had six levels of English. We were now a separate Adult language center school with an enrollment of about 600 students. Students still attended voluntarily. We gave no grades. Teachers chose their own text books and decided when a student was ready for a higher level at any time during the term. Students still entered our classes at any time during the term, but our school became so popular that we had waiting lists of people wanting to enter our overfilled classes.

Using the pretests I developed and with the help of teachers in the school who helped grade and administer tests, we began to set norms for new students by testing students in already existing classes of more or less homogeneous abilities. We still had some problems, it

was easy to place a new student who spoke no English at all or who knew a lot of vocabulary but no structures. If a student spoke some English, however, which level test should we give the student? Where should we give the test? Who would grade and assign the students to classes? We still had no counselors, no large rooms available, and no way of maintaining test security. Federal funds had been cut and I no longer had any released time.

In the meantime, other schools had become interested in our program and wanted to use the tests in their schools.

Mr. Roy Steeves of the California State Department of Education had been interested in our tests and with George Johnson of the San Francisco Unified School District obtained a federal grant for the district from the Southwestern Cooperative Educational Laboratory in Albuquerque, New Mexico to: 1. Make a short standardized placement test of two forms which would place students into the first three levels of English-Second-Language classes, 2. Set State norms on the new standardized test, and 3. Finish 400-500-600 pretests.

Results of the work done under that grant are as follows:

1. Standardized Test - from the 180 pretest items, Mrs. Jeanette Best and I selected 100 items. We gave this new test to 263 students at Alemany Adult School in January, 1969. From this test Mr. Richard Reyes and his assistant, Mr. Fred Gillette, from San Francisco State College made two tests of 50 questions each which we now call EPT 100-200-300 Forms A and B. Dr. Roger Cummings from San Jose State College supervised all statistical work. Mr. Reyes report (attachment A) provides further statistical information about the test.

2. Norms

In selecting places to give the test, we wrote letters to a number of school districts asking them to volunteer to be part of the testing program. Since drop out rates are high in adult programs and since students enter and leave all during the term,we needed an area with a large number of students in the program. We also needed to test groups using English-Second-Language learning materials rather than literacy or other first language materials.

Mrs. Jean Jacobs from Fremont Adult School in Sacramento and Dr. George Jensen from the Imperial Valley School District in El Centro volunteered for the program. Fremont Adult School has both day and night school heterogeneous classes. Students come from many language backgrounds. The Imperial Valley has only night school classes with almost all of the students speaking Spanish as a first language. Professor Alice Benz of the American Language Institute at San Francisco State College and Professor Dorothy Danielson from the Foreign-Born College English program at San Francisco State College also gave the test to their students. (See attachment C for a comparison of mean scores of all groups tested.)

Since school districts refer to their classes in different ways we grouped the classes by the textbook used in the class and referred to the levels as 100-200-300 to avoid confusion with grade levels referred to by first language teaching methods.

We were not surprised to learn that our test was too easy for the college groups, but we did not anticipate that except for Calexico one of the schools in the Imperial Valley, all regular level 100 classes would fall into a range of means from 8.11 - 20.00, that all level 200 classes would fall into a range of means from 20.13 to

27.95, and that all level 300 classes would fall into a range from 29.80 - 37.13.

We tested in February, the beginning of a term and again in May, the end of the term because we hoped to do a gains study. We wanted to determine the effects of different school environments on language learning ability and to investigate the feasibility of using EPT 100-200-300 as a measure of achievement over the period of one semester. Such a study can be made on the students at Alemany Adult School, but because a large number of students in the migrant worker Imperial Valley area had left school before the end of the term to follow the lettuce crop in another area and students at Sacramento had taken jobs for the summer, the sampling of students in those areas taking tests both in February and in May are too small for a reliable study. A pilot gains study revealed, however, that EPT 100-200-300 can be used as an achievement test over a period of a semester. Significant gains were made by all levels in all geographical areas. (See Attachment C.)

3. 400-500-600 pretests

   By June, 1969, Mrs. Jeanette Best and Mrs. Virginia Biagi at Alemany finished the pretests for levels five and six with reliabilities ranging from .54 to .87 and form correlations ranging from .77 to .85. They also tested other adult schools in our area and obtained norms for each level. However, SWCEL decided not to renew the grant to put these texts into a booklet which would place the students in the last three levels of English (400-500 → 600). (See Attachment D for pretest statistics.)

Coming full circle, our last efforts have been voluntary in making two forms of an experimental test for the last three levels. We call this test EPT 400-500-600 Forms X & Z. We do not have the excellent help of Mr. Reyes to choose the items and we must rely again upon schools having upper level classes and people trained in testing and handling test security volunteering to give Forms X & Z and set norms. Fortunately the American Language Institute, San Francisco State College English Program for Foreign Born Students, and City College in San Francisco, as well as Cambria Adult School in Los Angeles will test their students in February, 1970 and send us the results so that we can revise and refine our upper level test.

Another problem we have on both tests is protecting the security. There are some in the adult education field who feel that since part of the work was done under a Federal Grant, the test should be available to anyone interested in looking at it or using it with no arrangements for protecting test security. Some feel it should not be restricted in any way and should be placed like a book or article on a shelf in a reference library.

We also have a problem checking the validity of the tests because we do not have any outside form of comparison. Teachers in Adult Education classes do not give or keep grades for their students. Many teachers send all students on to the next level whether they have passed or failed the work. On pretests for levels 1 and 2, we asked teachers to group students in the top 25%, middle 50% and bottom 25% before the students took the tests. A t test showed differences were in the expected direction and probability was in every case better than .01. We also plan to compare our upper level student scores on current standardized tests for first-language learners and on tests designed for College Foreign born students.

We also have a test administration problem in our school. We have found that some teachers are unable to give a standardized test and maintain necessary test security without special training. We also have no place or personnel available to test new students entering our program after the classes are established. At the present time, we give the tests at the end of a term to all of our students to see if they are ready for the next level. Many new students take the tests at this time. A teacher's judgment obviously comes first and a few regular students are sent on to a higher level even if scores are not high enough because we feel it is important for a student to feel he is learning and successful.

A teacher committee is trying to work out a plan where all new students would be assigned quickly to a class on a temporary basis. Then one day every week, one teacher could take a double class while the released teacher (trained in test administration) gave the test to the new students.

An example of my Fall 1969 class which meets from 10:00 a.m. to 12:00 p.m. five days a week reflects a problem we now have after a term begins. While my average daily attendance was around 26, and my active enrollment averaged around 35, 79 students passed through this class during the semester and 28 remained at the end. This was not a loss of 51 students since 21 of these remained in the school after being transferred up or down one or two times until they reached a level where they could work successfully and still be challenged. Seven were referred to other schools and colleges, 3 moved out of the city, 4 found jobs, 6 were ill or had baby sitting problems. I don't know what happened to 10 of the students. Perhaps if they had had a better placement at the beginning, they might have stayed in the school.

One of the major problems is that in addition to a placement test, teachers need diagnostic tests that will tell them where their students need work. Professor Da..ielson at San Francisco State College is working with a number of Master Degree candidates who hope to come up with some diagnostic tests. My placement test will not serve as a diagnostic test. It also has no value for placing students in literacy classes. The test, however, has been used in placing junior high students in bilingual ESL programs.

Hopefully the work we have done at Alemany Adult School in San Francisco will be helpful to others interested in placing students at structural levels they can handle, but still find challenging. Hopefully someone will develop diagnostic tests in all areas, production as well as reception.

In summary, while we still have many problems in administering and using our placement tests, we have one placement test for our lower levels that has been standardized and which has a high reliability (as well as an experimental form for the upper levels) that place our students in classes better than previous methods. We still have to move a student or two, but not the large numbers of students we did before. The morale in the school is better. In short, both students and teachers are happier when placements are made more accurately.

Attachment A

TESTING REPORT by RICHARD REYES
on
ENGLISH-SECOND-LANGUAGE PLACEMENT TEST for LEVELS 100-200-300

Purpose:

The purpose of this study was to develop two equivalent forms of an English placement exam design, specifically for non-English speaking subjects.

The exam will be oriented entirely toward grammar placement and will utilize knowledge of written grammar as the differentiation factor in placement.

The exam will consist of a maximum of 50 items with a half-hour time limit. All questions will be of objective nature with three alternatives.

The exam is designed to place students in levels 1 - 3 out of a possible six levels. A similar exam will be developed to place students in levels 4 - 6.

Test items will come from a pool of 180 items previously developed and validated by the Alemany Adult Education School by Donna Ilyin.

METHOD (pretest):

From past statistical work, a total of 110 questions were selected from the original pool of 180 items developed at Alemany School. The basic rules for selection of items were:

1) Equal number of items from each of the three levels involved (1 - 3 levels).
2) Item difficulty should be in the range of 50 to 80 percent.
3) Item reliability should be greater than .45.
4) Items should be equal as far as representing an area of grammar necessary to mastering of the English language.

The 110 items were then split into two forms using the rules governing the selection of particular items as much as possible when applicable.

These two forms were then administered to a total of 263 subjects.

The purpose of this administration was to determine the equivalence of the two forms and to determine the reliability of each form. The 263 subjects were equally distributed within levels 1 - 3.

The results of the test administration were as follows:

1) Correlation between form A and B was .93.
2) Reliability of form A was .92.
3) Reliability of form B was .92.
4) Overall reliability of forms A and B scored together was .96.

As indicated earlier a maximum of 50 items was desirable due to time limitation in administration. To accomplish this, five items from each form of the exam were eliminated using the same rules as those used in arriving at the original figure of 55. After this was accomplished the two forms of the exam were rescored. The results of the rescoring indicated little change in the original statistics. The results were as follows:

1) Correlation between forms - .93.
2) Reliability of form A - .91.
3) Reliability of form B - .92.
4) Overall reliability of forms A and B - .95.

In order to assure that the two forms were equivalent at each of the levels a product moment correlation was obtained for each level between forms A and B. The results were as follows:

1) Level 1 - R=.93
2) Level 2 - R=.92
3) Level 3 - R=.90

From the results, it was concluded that the two forms were in fact equivalent at each of the desired levels.

Reliabilities at each of the levels were not computed due to the reduced number of students, however, overall reliability of forms A and B were sufficiently high enough (.91 and .92) to conclude that the 50 item tests were reliable. These reliabilities will be computed on the initial administration of each form.

METHOD (First Examination):

In order for this examination to be of value, a norm group was necessary. In an attempt to establish these norms for the examination, a total of 243 students were tested at Alemany School.**Students were randomly assigned forms and given a total of a half hour to complete the examination.

Results of the administration are shown in the following table:

N		MEAN	SD	SEM	RELIABILITY
33	Level 3	33.79	6.09	2.88	.78
44	Level 2	27.14	7.73	3.05	.84
44	Level 1	17.66	7.59	2.97	.85
41	Level 3	34.46	7.03	2.85	.84
44	Level 2	26.48	6.79	3.09	.79
37	Level 1	19.03	7.48	3.04	.84
121	Form A	25.5	9.74	3.02	.90
122	Form B	26.9	9.40	3.06	.89

Results of this administration support the results of the pretest data. Reliability ranged from a low of .78 to a high of .85. Forms A and B produced reliabilities of .90 and .89 respectively. The reason for the lower level reliability at each of the different levels is due in part to the reduced number of students in each group.

**Alemany Adult School has an active enrollment of about 600 foreign students per semester from about sixty different countries. In May, 1969 the median age was 24.2, but the range was from 17 to 80 years. More than half of these students had not finished the 12th year in their own countries, but 197 had college training and 22 were college graduates.

About half of the students had little or no English instruction before coming to the United States, and many of the others had only a workbook type instruction leaving them with little ability for reading comprehension or to speak or understand the simplest conversations. Most of the students lived with families or relatives who probably spoke their native language.

Means and standard deviations were basically similar at each of the levels. Level one shows variation, but this was not significant and is due in part to the limited sample in this level. The combining of the levels also failed to influence the mean difference between forms A and B. A test of significance indicated no difference between the means of the two forms.

From the previous page results, it was concluded that both forms could be combined to form one percentile range. This was done mainly for the convenience of teachers using the examination. The percentile range is as follows:

ALL LEVELS COMBINED FORMS ALBANY SCHOOL PERCENTILES COMBINED

SCORE	FREQUENCY	CUMULATIVE FREQUENCY	PERCENTILE
50	0	243	99
49	2	243	99
48	1	241	99
47	1	240	99
46	0	239	98
45	2	239	98
44	2	237	98
43	1	235	97
42	1	234	96
41	4	233	95
40	5	229	94
39	6	224	92
38	2	218	90
37	11	216	89
36	7	205	84
35	6	198	81
34	10	192	79
33	9	182	75
32	9	173	71
31	10	164	67
30	5	154	63
29	5	149	61

ALL LEVELS COMBINED FORMS — ALLEGANY SCHOOL — PERCENTILES COMBINED

SCORE	FREQUENCY	CUMULATIVE FREQUENCY	PERCENTILE
28	8	144	59
27	4	136	56
26	11	132	54
25	10	121	50
24	12	111	46
23	7	99	41
22	9	92	38
21	9	83	34
20	7	74	30
19	11	67	28
18	11	56	23
17	8	45	19
16	6	37	15
15	7	31	13
14	5	24	10
13	4	19	8
12	0	15	6
11	2	15	6
10	1	13	5
9	2	12	5
8	2	10	4
7	1	8	3
6	1	7	3
5	2	6	3
4	1	4	2
3	1	3	1
2	0	2	1
1	2	2	1

Attachment B

BOOKS USED IN E.S.L. 100, 200, AND 300

E.S.L. 100

Bernardo, Leo U. & Pantell, Dora F., English Your New Language, Book 1, Silver-Burdett and Company, 1967.

Breckenridge, Robert, Book I, publishing pending with McGraw Hill.

Dixson, Robert J., Modern American English, Book 1, Regents Publishing, 1962.

Dixson, Robert J., Regents English Workbook, Regents Publishing, New York, 1956.

English Language Services, English 900, Book 1 & 2, The MacMillan Company, New York, 1964.

English Language Services, English This Way, Books 1 and 2, The Macmillan Company, New York, 1963.

English Language Services, Intensive Course in English, Elementary, Part 1, English Language Services, 1963.

King, Harold V. and Campbell, Russel N., Modern English Primer, Part I, English Language Services, Inc., 1956.

National Council Teachers of English, English for Today, Book I McGraw Hill, 1962.

E.S.L. 200

Dixson, Robert J., Modern American English, Book 2, Regents Publishing, 1962.

English Language Services, English 900, Books 2 & 3, The MacMillan Company, 1964.

English Language Services, English This Way, Books 3 & 4, The MacMillan Company, 1963.

English Language Services, Intensive Course in English, Elementary, Part 2, English Language Service, Inc., 1963.

Hall, J., Practical Conversation in English Intermediate, Regents Publishing Company, 1967.

National Council of Teachers of English, English for Today, Book 2, McGraw Hill, 1962.

Wright, Audrey L. and McGillivray, James H., Let's Learn English, American Book Company, 1960.

Fries, American English Series, Book 3, D.C. Heath and Co., Boston, 1953.

E.S.L. 300

Bernardo, Leo U. & Pantell, Dora F., English Your New Language, Book 2, Silver-Burdgett and Company, 1967.

Binner, Vinal O., American Folktales I, Thomas Crowell Co., 1968.

Breckenridge, Robert, Breckenridge Book II, publishing pending with McGraw Hill.

Dixson, Robert J., Complete Course in English, Book I, Regents Publishing, 1955.

Dixson, Robert J., Modern American English, Book 3, Regents Publishing, 1962.

English Language Services, English 900, Book 4, The MacMillan Company, 1964.

Franklin, Harry B., Meikle, Herbert G., Strain, Jeris E., Vocabulary in Context, University of Michigan Press, 1964.

Lado, Robert & Fries, Charles C., English Pronunciation & English Sentence Patterns, University of Michigan Press, 1954.

Taylor, Grant, Learning American English, McGraw Hill Book Company, 1956.

Wilcox, Lillian A.; Thomas, Lydia A., Reader's Digest Skill Builder # 3, Part 2, Reader's Digest Services, Inc., 1959.

Wright, Audrey, Barrett, Ralph P., Van Syoc, W. Bryce, Let's Learn English, Book 4, American Book Co., 1968.

ATTACHMENT C

MEANS AND COMPARISON OF MEANS

NO FORMULA USED FOR REGRESSION EFFECTS

(Raw means only)

Day Schools:

	Feb. Means	Sec.	City	Group I No. Feb.	Group 2 No. Taking Both Tests	Pre-Test Feb. Means Group 2	Post-Test May Means Group 2	Mean Diff.	T-Value
Level I	26.26	A	SF	26	15	28.60	36.06	7.46	6.05*
	18.18	B	SF	22	19	19.10	24.89	5.78	
	20.00	C	SF	37	13	18.61	25.00	6.38	
	7.16	D	SF	24	8	13.12	21.87	6.31	
	16.69	A	Sacto	13	40	17.75	24.32	6.57	7.73*
			(night	14)	17	15.29	21.00	5.70	5.06*
			(night	21)					
Level II	33.71	A	SF	28	20	35.30	38.35	3.05	3.25*
	27.65	B	SF	31	21	28.33	33.09	4.76	
	25.45	C	SF	22	16	25.93	28.50	2.56	
	24.62	A	Sacto	24	37	27.29	31.10	3.81	5.71*
			(night	14)	17	23.52	27.52	4.00	2.30**
			(night	19)					
Level III	34.59	A	SF	37	28	35.07	36.17	1.10	
	30.76	B	SF	30	19	30.36	36.00	5.63	
	29.80	A	Sacto	26	47	33.17	36.10	2.93	2.42**
		B	Sacto	23	33	30.93	34.48	3.54	4.01*
			(night	23)					
ALI	41.67	SF State	SF	40	33	40.63	46.21	5.58	
College English for Foreign Students	45.25	SF State	SF	81	Scored too high to retest.				

Notes:

(1) San Francisco 1A & 2A classes are atypical accelerated classes and ID is very slow.

(2) Sacramento Day and Night Schools are taken together by statistician.

(3) ALI students did not take alternate forms.

* Significant to .01 level

** Significant to .05 level

ATTACHMENT C - continued

MEANS AND COMPARISON OF MEANS

NO FORMULA USED FOR REGRESSION EFFECTS
(Raw means only)

Night Schools:

	Feb. Means	Sec.	City	Group I No.Feb.	Group 2 No.Taking Both Tests	Pre-Test Feb.Means Group 2	Post-Test May Means Group 2	Mean T-Diff.Val.
Level I	15.16**	Washington	SF	28	See Below			
	8.11	A	Sacto	14	See Day Schools			
	18.52	B	Sacto	21				
	12.57	Calexico		20	5	14.00	20.20	6.20
	19.55	El Centro		28	9	13.77	18.11	4.34
	11.35	Brawley		19	10	9.99	14.30	4.31
Level I&II	17.11	Washington	El Centro	20	4	14.75	15.00	.25
Level II	27.95**	Washington	SF	20	See Below			
	23.21	A	Sacto	14	See Day Schools			
	27.36	B	Sacto	19				
	14.81*	Calexico		17	See Below			
	25.33	El Centro		9	4	21.50	26.50	5.00
					6	17.83	16.17	-1.66
Level II&III	20.13	Brawley		29	12	23.92	27.50	3.58
Level III	35.58**	Washington	SF	17	See Below			
	35.73	Sacramento		23	See Day Schools			
	25.16*	A	Calexico	12	See Below			
	20.94	B	Calexico	17	2	36.00	26.00	-10.00
	33.00	Citizen	El Centro	20	Class received citizenship a week before retest.			
				29	13	21.08	24.85	3.77
				65	15	32.13	34.33	2.20
					63	18.42	21.41	2.98 3.88

*Calexico combined 2 and 3 and dropped a class.

**Combined 1, 2, 3 Washington Night SF

Gains on all Imperial Valley Students levels I, II, & III
T-Value 3.88 is significant to .01 level.

ATTACHMENT D

STATISTICS FOR ENGLISH PLACEMENT TEST PRETESTS

I Level 100 Tests

A. Forms L & M

Sample: 100 Level I students at the end of the semester from various language backgrounds who attended Alemany Adult School in May, 1967. The term began February 2, 1967.

	Form L	Form M
Maximum possible raw score	30	30
Mean	16.47	16.41
Standard deviation	5.96	6.03
Reliability using Kuder, formula 20	.85	.85
Semi-interquartile range	2.33	2.35

Pearson Product - Moment correlation between Forms L & M = .80

Validity: Since no grades are given to students enrolled in Alemany Adult School no data was available of the type usually used in validity studies. To make the rating problem as simple as possible for the teachers, they were asked to assign students to one of three groups - HIGH (consisting of the top 25 percent), MID (consisting of the middle 50 percent), and LOW (consisting of the low 25 percent). In order to control for bias, teachers were asked to make these ratings before they had seen the test scores for their students.

Tests of significance between the mean scores for each of these three groups were computed by means of t scores. Differences were all in the expected direction and probability was in every case better than .01.

T TEST RESULTS

	FORM L			FORM M		
	Mean	Sigma	t	Mean	Sigma	t
Hi	22.23	5.30		23.55	3.01	
Mid	16.33	3.73		15.95	4.20	
Low	12.76	5.09		12.52	4.92	
Hi/Mid			4.57			8.2
Hi/Low			8.23			9.43
Mid/Low			3.00			2.86

B. Forms A, B, C, & D

Teachers analyzed a list of 90 suggested test items prepared by the test developer. They made suggestions and additions. A trial test was made consisting of 82 items and given to 124 students in the beginning levels.

A comparison of items missed by the top 27% (32 students) with the 27% low (32 students) yielded a list of discriminating items.

Two experimental placement tests were made from the 60 best discriminating items and called Forms C and D. (Two other Forms A and B were discarded as unsatisfactory.)

In October, 1966, a Pearson Product-Moment correlation between Forms C and D on a sample of 86 Level I students was .69 and on a sample of 86 Level II students was .72.

Sample: 57 Level II students at the beginning of the semester from various language backgrounds attending Alemany School in March, 1967.

	Form C	Form D
Maximum possible raw score	30	30
Mean	21.34	19.88
Standard deviation	5.01	4.95
Reliability	.82	.78
Semi-interquartile range	2.12	2.31

Validity: Differences in mean scores between Hi, Mid and Low groups were mainly in the expected direction. Except for the differences in the t test between the Mid/Low for Form D, probability was better than .05 in one case and better than .01 in the remaining.

T TEST RESULTS

	FORM C			FORM D		
	Mean	Sigma	t	Mean	Sigma	t
Hi	24.08	1.25		23.73	4.24	
Mid	22.00	4.95		20.26	4.24	
Low	18.64	4.16		17.36	5.39	
Hi/Mid			5.94			2.22
			4.19			3.08
			2.62			1.57

II Level 200 Tests

A. Forms N & O

Sample: 53 Level II students at the end of the semester from various language backgrounds who attended John Adams Annex in May, 1967. The term began February 2, 1967.

	Form N	Form O
Maximum possible raw score	30	30
Mean	18.04	17.11
Standard deviation	4.63	4.78
Reliability using Kudar formula 20	.74	.75
Semi-interquartile range	2.35	2.38

Pearson Product- Moment correlation between Forms N & O = .86

NOTE: The administration of the test was given under less than standard conditions. It was not adequately proctored due to a misunderstanding on the part of the proctors, however, the results are not inconsistent with previous forms. (See the statistics for part II Forms A and B, page 4)

Validity: Differences in mean scores between Hi, Mid and Low groups were mainly in the expected direction. Except for the difference in the comparison of Hi means with Mid means, probability was better than .01.

T TEST RESULTS

	FORM N			FORM O		
	Mean	Sigma	t	Mean	Sigma	t
Hi	20.25	4.78		19.25	4.88	
Mid	18.64	3.89		17.36	4.86	
Low	14.00	3.81		13.27	3.54	
Hi/Mid			1.00			1.08
Hi/Low			3.47			3.38
Mid/Low			3.27			2.74

B. Forms A, B, C, & D
Suggested items analyzed by teachers = 189 for pretests.

Trial Placement Tests	Part I	Part II
Items	87	87
Sample: # Students	177	110
Number of 27% high scores analyzed	49	30
Number of 27% low scores analyzed	49	30

Four experimental placement tests were then made each containing 30 three choice items and called Forms A, B, C, & D.

Reliability correlations (r) between Forms A, B, C, & D and the sample number of students (N) in October, 1966 were as follows:

	Level II	Level III
Form A/B	r = .83 N = 61	r = .64 N = 33
Form C/D	r = .65 N = 59	

	Level II	Level III
Form A/C	r = .62 N = 73	
Form A/D	r = .61 N = 42	
Form B/C	r = .61 N = 41	r = .79 N = 14
Form B/D	r = .58 N = 44	r = .66 N = 10

Sample: 50 Level III students at the beginning of the semester from various language backgrounds who attended Alemany School in March, 1967.

	Form A	Form B
Maximum possible raw score	30	30
Mean	20.18	20.54
Standard deviation	5.06	5.07
Reliability	.81	.82
Semi-interquartile range	2.22	2.17

Validity: Due to the practice of grouping students into ability sections which are not always consistant throughout the day, at Level III, only one teacher's ranking of students per class was used for the t tests. The differences tended in the expected direction. Except for the t test for the difference of means between the Mid/Low group in Form A and the Mid/Hi group in Form B, probability ranged from .05 to better than .01.

T TEST RESULTS

	FORM A			FORM B		
	Mean	Sigma	t	Mean	Sigma	t
Hi	24.00	4.21		23.71	5.99	
Mid	19.39	5.00		19.56	5.12	
Low	15.33	4.27		15.33	3.01	
Hi/Mid			2.31			1.62
Hi/Low			4.03			3.25
Mid/Low			1.93			2.46

III No standardization work or form correlations were performed on Level 300 tests, but teachers analyzed 280 questions.

Trial Placement Test	Part I	Part II
Items	75	75
Sample: # Students	67	67
Number 27% high scores	18	18
Number 27% low scores	18	18

Placement test Forms A and B were made for this level but not standardized.

IV Level 400 Tests

A. Forms R & S

1. Sample: 51 Level IV students at Alemany tested at the end of the term June, 1969.

	Form R	Form S
Maximum possible raw score	30	30
Mean	20.06	19.06
Standard deviation	4.53	4.62
Reliability	.74	.75
Semi-interquartile range	2.31	2.32
Pearson Product - Moment correlation between R & S	.66	

2. Sample: 23 Level V students at Pacific Heights tested at the end of the term June, 1969.

	Form R	Form S
Maximum possible raw score	30	30
Mean	19.57	18.00
Standard deviation	5.04	4.86
Reliability	.79	.77
Semi-interquartile range	2.29	2.34
Pearson Product - Moment correlation bewteen R & S	.68	

B. Forms A, B, C, & D

Trial Placement Tests A, B, C were made from a list of 46 suggested items and four variations of four of the items.

1. Sample: 60 Level IV students at the end of the term May, 1967.

Items	50
Number 27% high scores analyzed	16
Number 27% low scores analyzed	16

Forms A and B were exactly alike except for a typing variation. Form C is a slight variation of B because we did not have enough discriminating items. In May, 1967 reliability correlations between Forms B and C on a sample of 27 IV A students was .87 and on a sample of 23 IV B students was .76.

2. Sample: 52 Level IV students at the end of the semester from various language backgrounds attending Alemany Adult School in May, 1967.

	Form B	Form C
Maximum possible raw score	30	30
Mean	19.76	19.90
Standard deviation	5.14	5.26
Reliability	.82	.82

	Form B	Form C
Semi-interquartile range	2.20	2.26

Validity: Because of ability grouping practices which are not consistent throughout the day, only one teacher's ranking of students per class was used for the t tests. The difference of mean scores between Hi, Mid and Low were in the expected direction and probability was better then .01.

T. TEST RESULTS

	FORM B			FORM C		
	Mean	Sigma	t	Mean	Sigma	t
Hi	25.83	2.30		25.17	2.07	
Mid	19.00	4.08		19.48	4.84	
Low	15.17	4.31		16.58	3.47	
Hi/Mid			6.57			5.13
Hi/Low			7.56			7.34
			2.59			2.12

Suggested items were taken from lower pretest items too difficult for previous levels and used for Form D

1. Sample: 80 Level IV students at Alemany tested at the end of the term May, 1968.

Trial Placement Test - Spring, 1968

Items	75
Number 27% high scores analyzed	21
Number 27% low scores analyzed	21

Placement Tests C (from 1967) and D (from Spring 1968)

2. Sample: 72 Level IV students at Alemany tested at the end of the term May, 1968.

	Form C	Form D
Maximum possible raw score	30	30
Mean	19.00	20.88
Standard deviation	4.66	5.66
Reliability	.75	.85
Semi-interquartile range	2.31	2.22

Pearson Product - Moment correlation between C and D: None provided by statistician

V. Level 500 Tests

A. Placement Tests T & U

1. Sample: 34 Level V students at Alemany tested at the end of the term June, 1969.

	Form T	Form U
Maximum possible raw score	30	30
Mean	17.56	18.59
Standard deviation	3.64	4.94
Reliability	.54	.78
Semi-interquartile range	2.47	2.34
Pearson Product - Moment correlation between T and U	.77	

2. Sample: 32 Level VI students at Pacific Heights tested at the end of the term June, 1969.

	Form T	Form U
Maximum possible raw score	30	30
Mean	20.25	20.63
Standard deviation	5.32	4.42
Reliability	.81	.75
Semi-interquartile range	2.32	2.26
Pearson Product-Moment correlation between T and U	.82	

B. Placement Tests A & B

Suggested items were taken from a compilation of about 200 questions.

1. Sample: 50 Level VI students at Alemany tested at the beginning of the term February, 1969.

Trial Placement Tests - the beginning of the term February, 1969

Items	75
Number 27% high scores analyzed	14
Number 27% low scores analyzed	14

2. Sample: 34 Level VI students at Alemany tested in the middle of the term March, 1969.

	Form A	Form B
Maximum possible raw score	30	30
Mean	20.85	20.89
Standard deviation	3.81	4.15
Reliability	.63	.75
Semi-interquartile range	2.32	2.07
Pearson Product - Moment correlation between A and B	None (because statistician did not do correlation on the same students)	

VI Level 600 Tests

A. Forms A & B

Suggested items taken from compilation of about 200 questions.

1. Sample: 70 Level VI students at Alemany tested beginning of term Spring, 1969.

Trial Placement Tests - beginning of Spring term 1969

Items	75
Number 27% high scores analyzed	20
Number 27% low scores analyzed	20

2. Sample: 37 Level VI students at Alemany tested at the end of the term June, 1969.

Placement Tests A and B

	Form A	Form B
Maximum possible raw score	30	30
Mean	21.92	21.00
Standard deviation	3.93	3.83
Reliability	.68	.65
Semi-interquartile range	2.21	2.28
Pearson Product - Moment correlation between A and B	.80	

3. Sample: 13 Civics students at Pacific Heights tested at the end of the term June, 1969.

	Form A	Form B
Maximum possible raw score	30	30
Mean	18.23	19.15
Standard deviation	6.15	5.25
Reliability	.87	.82
Semi-interquartile range	2.21	2.24
Pearson Product - Moment correlation between A and B	.85	

ERIC Clearinghouse
FEB 25 1970
on Adult Education